

## COMPUTED TOMOGRAPHY

All skulls ( $N = 36$  carnivore species) were scanned using a General Electric Lightspeed 4 slice CT scanner in the Department of Radiology at Michigan State University. Each skull was aligned in the scanner rostrocaudally in order to replicate the natural anatomical position of the head. Parameters for each scan were as follows: 0.625 mm slice thickness, 30 cm field of view, 5.62 mm/rotation table speed and 0.562:1 pitch. CT images were saved in DICOM (Digital Imaging and Communications in Medicine) Centricity (version 2.2) format. Virtual three-dimensional endocasts were created using the software package MIMICS 11.02 (Materialise, Inc., Ann Arbor, MI, U.S.A.), as described previously (Sakai, Arsznov, Lundrigan, & Holekamp, 2011a, 2011b). Assessments of total endocranial and regional volumes were obtained using the MIMICS three-dimensional volume measurement operation. Total endocranial volume was defined as the volume extending from the rostral tip of the olfactory bulbs caudally to the foramen magnum. Further details are available in Sakai et al. (2011a, 2011b).

## RELATIONSHIP BETWEEN SOCIABILITY AND FITNESS

### *Methods*

#### *Hourly joining rate*

We observed hyaenas from vehicles daily between 0530 and 0930 hours and between 1700 and 2000 hours. We initiated observation sessions whenever we saw a hyaena or group of hyaenas separated from conspecifics by at least 200 m. Upon initiating an observation session, we conducted a behavioural scan that identified each hyaena; these scans were repeated every 20 min. During each session, we recorded whenever a hyaena entered or left a session in progress using all-occurrence sampling (Altmann, 1974). To measure sociability, we calculated the hourly rate at which each adult female hyaena joined other groups of hyaenas when she was initially alone during the period between 1996 and 2002. We only included hyaenas as focal subjects that were observed for at least 20 h during that period. A hyaena was determined to have joined

others if it approached and remained within 200 m of at least one other hyaena rather than just passing through and moving on. The original (joined) group could either be stationary or travelling together. To calculate the hourly group-joining rate for each female, we divided the number of joining events for each individual during the 7-year study period by the number of hours in which the individual was observed during that same period.

### *Fitness measures*

As indicators of reproductive success, we used the total number of cubs borne by each female over her lifetime, as well as the proportion of her cubs that survived to three life-history milestones: den independence, weaning, and 24 months of age. We created a set of generalized linear and second-order polynomial models with these fitness components as the response variables, and group-joining rate, the maternal rank of the female and an interaction term as predictors. Only female hyaenas that were born after our study began and that died before our study ended were included in these reproductive success analyses. We also used hyaenas' longevity as a response variable to determine whether sociability is related to survival. For this data set, we used the same criteria for inclusion, except we included only females for which we had known death dates. We created both generalized linear and second-order polynomials with longevity (in months) as the response variable. As fixed effects, we included the group-joining rate and the average lifetime standardized rank of the female, as well as an interaction term between these variables.

### *Statistical analysis*

We used the reproductive success and longevity of female hyaenas as response variables to determine whether sociability is related to fitness. We used the group-joining rate (defined above) of each female as the measure of sociability in this analysis. We calculated a mean social rank for each female over her lifetime, and then standardized these mean ranks. Only females for which we had known birth and death dates that were observed as adults for at least 15 h between 1996 and 2002 were included in this analysis. We adopted a Bayesian approach and created models with the Markov chain Monte Carlo for generalized linear mixed models (MCMCglmm)

analysis tool in R (Hadfield, 2010). We used a Poisson error structure and a relatively uninformative prior (Hadfield, 2010); however, using a more informative prior did not affect our results (Shaw, 2012). As fixed effects in each model, we included the group-joining rate and the average lifetime standardized rank of the female, as well as an interaction term between these variables. We tested fixed effects for significance by determining whether or not the 95% credible interval for each fixed effect included zero; variables for which the credible interval included zero were not significant predictors of the response variable, whereas those that did not include zero were considered significant. Our estimates are reported as posterior means with 95% credible intervals. For longevity and all measures of reproductive success, we used likelihood ratio tests to determine whether second-order polynomial models were able to predict these fitness measures better than linear models. All variables in these analyses were Z transformed.

## *Results*

There were 31 females that satisfied the criteria for inclusion in our analysis of the relationship between group-joining rate and reproductive success. In generalized mixed models, group-joining rate was not significant in predicting any measure of reproductive success. However, we found several significant interactions between group-joining rate and maternal rank. Although high-ranking hyaenas had consistently high reproductive success, low-ranking females with high group-joining rates enjoyed better reproductive success than did low-ranking females with low group-joining rates. The interaction between maternal rank and group-joining rate was significant for all measures of reproductive success associated with offspring survivorship, including the percentage of cubs surviving to weaning ( $0.022 \pm 0.007$ ,  $N = 31$ ; generalized linear model:  $t_{27} = 3.292$ ,  $P = 0.003$ ), the percentage of cubs surviving to den independence ( $0.024 \pm 0.007$ ,  $N = 31$ ; generalized linear model:  $t_{27} = 3.447$ ,  $P = 0.002$ ) and the percentage of cubs surviving to reproductive maturity ( $0.018 \pm 0.007$ ,  $N = 31$ ; generalized linear model:  $t_{27} = 2.615$ ,  $P = 0.014$ ). There were 34 female hyaenas that satisfied the criteria for inclusion in our analysis of the relationship between group-joining rate and longevity. The age at

death of these females averaged  $112.38 \pm 1.30$  months. Group-joining rate significantly predicted age at death; hyaenas that joined groups more often lived longer than did those that joined groups less often ( $19.820 \pm 7.354$ ,  $N = 34$ ; generalized linear model:  $t_{31} = 2.695$ ,  $P = 0.011$ ). No interaction terms containing group-joining rate significantly predicted longevity, and second-order polynomial terms were not significantly better than linear models at predicting age at death ( $P > 0.1$ ).

## ZOO STUDY

### *Methods and Results*

Full methodological details are available in Benson-Amram, Dantzer, Stricker, and Holekamp (n.d.).

### *Administration of puzzle box trials in zoos*

We ran puzzle box trials with 146 individuals (from 39 mammalian carnivore species in nine families) maintained at nine North American zoos: St Louis Zoo, St Louis, MO; Bergen County Zoo, Paramus, NJ; Binder Park Zoo, Battle Creek, MI; Potter Park Zoo, Lansing, MI; Columbus Zoo, Columbus, OH; The Living Desert, Palm Desert, CA; Wild Canid Survival and Research Center, Eureka, MO; Turtle Back Zoo, West Orange, NJ; Denver Zoo, Denver, CO.

Both large and small puzzle boxes allowed subjects to see and smell the bait inside, and all worked in exactly the same way such that the animal had to slide a simple bolt latch sideways to have the hinged door swing open so the food inside could be accessed, or move the box around until it was oriented such that the bolt would fall open, as described previously (Benson-Amram & Holekamp, 2012; Benson-Amram, Weldele, & Holekamp, 2013). Baits were chosen based on what keepers told us was the favourite food of each individual tested; baits ranged from bamboo to dead baby goats, but in no case could the subjects obtain the bait from the box unless the door was opened.

All subjects (146 individuals from 39 species) were fasted for 24 h before testing, and all trials were videotaped in their entirety. Each subject was briefly moved to an adjacent enclosure while the baited box was placed in the animal's home enclosure, the latch handle was set to protrude at a 90° angle from the door, and a tripod-mounted video camera was aimed to centre on the box from just outside the home enclosure. Trials lasted 30 min or until the animal obtained the bait from the box.

We conducted 516 trials in total as we tested one to nine individuals per species (mean = 3.5). The total number of trials conducted varied among species (Benson-Amram et al., n.d.; mean = 13 trials/species). Each individual within a species was tested with the puzzle box approximately three times. If the subject did not succeed at opening the puzzle box by its third trial, we conducted no further trials. Subjects that opened the puzzle box within the first three trials were considered to be successful at the task.

#### *Extraction of behavioural data*

With the help of an undergraduate assistant, S.B.A. performed all data extraction from videotapes of zoo trials. Our measure of success was simply whether or not the subject was able to open the puzzle box during its trial. We extracted performance measures from video footage as described earlier (Benson-Amram & Holekamp, 2012; Benson-Amram et al., 2013). To score behavioural diversity, we looked for 13 behaviours that all subjects were physically capable of performing, and scored whether or not an individual showed each behaviour. Each individual thus received a score from 0 to 13. We only scored each subject's first trial with the puzzle box.

#### *Total and regional brain volumes*

We obtained total brain volume (in ml) and adult body mass (in kg) data for each tested carnivore species from previously published data sets (Finarelli & Flynn, 2009). We also obtained measures of total endocranial volume and the volume of the anterior cerebrum (AC), posterior cerebrum (PC), total cerebrum (PC+AC) and the cerebellum plus brainstem (Cb+B<sub>s</sub>) and body mass (in kg) from our previously published data set (Swanson, Holekamp, Lundrigan,

Arsznov, & Sakai, 2012).

### *Statistical methods*

To control for the shared evolutionary history among our subject species, we used phylogenetic generalized least squares (PGLS) models to examine the relationship between brain volume, regional brain volume and group size (a measure of social complexity) and a species' mean performance during puzzle box trials. A species' mean performance was defined as the number of successful individuals opening the box divided by the total number of individuals tested in that species. We used an updated version of a recent mammalian supertree phylogeny (Bininda-Emonds et al., 2007) from Fritz, Bininda-Emonds, and Purvis (2009) and pruned species not in our sample in R (version 3.0.2: R Core Team, 2014) using the package 'geiger' (version 2.0.3: Harmon, Weir, Brock, Glor, & Challenger, 2008). We used group size (from Swanson et al., 2012) as a proxy for social complexity because Swanson et al. (2012) found this was no more or less effective than using PC2 from a principal component analysis of several different measures of social complexity.

We used two different data sets, a larger one containing 39 tested species for analysis of effects of total brain volume on performance, and a smaller one containing the 17 tested species for which we also had brain region data from CT scans used to analyse effects of size of specific brain regions on performance (all data appear in Benson-Amram et al., n.d.). We regressed performance on mass-corrected total brain volume in a PGLS model that included average group size, and average group size<sup>2</sup> and the covariates work time, behavioural diversity and manual dexterity. We also regressed performance on ln-transformed total brain volume in separate PGLS models to determine whether uncorrected brain volume influenced a species performance (as in MacLean et al., 2014).

In our analyses of the relative effects of different brain regions on performance in the puzzle box test in the restricted data set ( $N = 17$  species), we first calculated the volume of the four different brain regions (PC, AC, cerebrum, Cb+Bs) relative to overall endocranial volume (as in Swanson et al., 2012). We then regressed the volume of each brain on the overall endocranial volume minus the volume of that specific region (described in Swanson et al., 2012).

We also determined mass-corrected total brain volume in this restricted data set ( $N = 17$  species) by regressing ln-transformed total brain volume on ln-transformed adult body mass. The residuals from these models were then used as measures of relative brain region volume or mass-corrected total brain volume.

We used an information-theoretic approach (Akaike's information criterion corrected for small sample sizes,  $AIC_c$ : Burnham & Anderson, 2002) to determine whether the relative volume of any specific brain region better predicted success in opening the puzzle box than mass-corrected total brain volume. We developed 10 candidate PGLS models to explain performance at opening the puzzle box that contained either (1) mass-corrected brain volume, (2) relative PC volume, (3) relative AC volume, (4) relative cerebrum volume, or (5) relative Cb+Bs volume. All of these models contained the same covariates (number of trials, group size, group size<sup>2</sup>, work time, behavioural diversity, manual dexterity) but differed regarding whether they contained mass-corrected total brain volume or the relative volume of a specific brain region (PC, AC, cerebrum, Cb+Bs).

## References

- Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour*, 49, 227–265.
- Benson-Amram, S., Dantzer, B., Stricker, G., & Holekamp, K. E. (n.d.). *The evolution of problem-solving abilities in mammalian carnivores*. Manuscript in preparation.
- Benson-Amram, S., & Holekamp, K. E. (2012). Innovative problem solving by wild spotted hyenas. *Proceedings of the Royal Society B: Biological Sciences*, 279, 4087–4095.
- Benson-Amram, S. R., Weldele, M. L., & Holekamp, K. E. (2013). A comparison of innovative problem-solving abilities between wild and captive spotted hyenas (*Crocuta crocuta*). *Animal Behaviour*, 85, 349–356.
- Bininda-Emonds, O. R. P., Cardillo, M., Jones, K. E., MacPhee, R. D. E., Beck, R. M. D., Grenyer, R., et al. (2007). The delayed rise of present-day mammals. *Nature*, 446, 507–512.

- Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multi-model inference: A practical information-theoretic approach*. New York, NY: Springer.
- Finarelli, J. A., & Flynn, J. J. (2009). Brain-size evolution and sociality in Carnivora. *Proceedings of the National Academy of Sciences of the United States of America*, 106, 9345–9349.
- Fritz, S. A., Bininda-Emonds, O. R. P., & Purvis, A. (2009). Geographical variation in predictors of mammalian extinction risk: big is bad but only in the tropics. *Ecology Letters*, 12, 538–549.
- Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software*, 33, 1–22.
- Harmon, L. J., Weir, J. T., Brock, C. D., Glor, R. E., & Challenger, W. (2008). Geiger: investigating evolutionary radiations. *Bioinformatics*, 24, 129–131.
- MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., et al. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences of the United States of America*, 111, E2140–E2148.
- R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Sakai, S. T., Arsznov, B. M., Lundrigan, B. L., & Holekamp, K. E. (2011a). Virtual endocasts: an application of computed tomography in the study of brain variations in hyenas. *Annals of the New York Academy of Sciences*, 1225, E160–E170.
- Sakai, S. T., Arsznov, B. M., Lundrigan, B. L., & Holekamp, K. E. (2011b). Brain size and social complexity: a computed tomography study in Hyaenidae. *Brain, Behavior and Evolution*, 77, 91–104.
- Shaw, K. C. (2012). *Variation among free-living spotted hyenas in three personality traits*. (Doctoral thesis). East Lansing, MI: Michigan State University.
- Swanson, E. M., Holekamp, K. E., Lundrigan, B. L., Arsznov, B., & Sakai, S. T. (2012). Multiple determinants of brain size and brain architecture in mammalian carnivores. *PLoS One*, 7, e38447. doi:10.1371/journal.pone.0038447.